## REMARKS

### A. Amendments to the Specification

The specification is amended to correct typographical errors and to address idiomatic issues. No new matter is added by the above amendments.

### B. Amendments to the Claims

Claims 1-5, and 7-23 are presented for examination. Claims 1, 2, 7, 8, and 10-13 are currently amended. Claim 6 is cancelled. Claims 14-23 are new.

Claim 5 was objected to under 37 CFR 1.75(c) as allegedly being in improper form because a multiple dependent claim cannot depend from any other multiple dependent claim. Applicants respectfully point out that claim 5 was previously amended in a Preliminary Amendment dated February 2, 2001, in which all multiple claim dependencies were removed. Thus, the current 37 CFR 1.75(c) rejection is inappropriate.

In a telephone interview with Examiner Blackwell, it was agreed that the current Office Action may have overlooked the aforementioned Preliminary Amendment dated February 2, 2001, and the current claim rejections may therefore not be based on the most current state of the claims. Nonetheless, it was agreed that Applicants should proceed with filing an Office Action Response with the understanding that a second, non-final Office Action may be appropriate in the future to give due consideration to the current claim language.

Claims 1-4, 6, and 10-13 were rejected under 35 U.S.C. §102(a) as being anticipated by Zamir et al. ("Web Document Clustering: A Feasibility Demonstration", ACM, August, 1998).

Additionally, claims 7-9 were rejected under 35 U.S.C. §103(a) as being unpatentable over Zamir.

The language of the original claims as filed had been amended in the Preliminary Amendment of February 2, 2001 to more clearly recite claim limitations. It would appear that examination of the claims as filed may have introduced undue confusion leading to the Examiner overlooking two aspects of

the present invention. The first aspect relates to the formula used in the present invention to calculate a relationship between clusters, and the second aspect relates to the use of the terms "AND" and "OR" in claim 7. Both aspects are address in turn below.

Firstly, claim 2 had stated that,

"the evaluation of relations among clusters under consideration as to whether they should be merged or not is performed on the basis of the number of documents commonly included in said clusters under consideration relative to the total number of documents included in said clusters under consideration"

Thus, the calculation of the relationship between clusters is based on the combined number (i.e. "total number") of documents in all the clusters under consideration. Zamir explains (in page 49, col. 1, third paragraph), that his determination of the relationship between clusters is determined on a cluster-by-cluster basis, and not based on a combination of clusters, as is required in claim 2. Specifically, Zamir explains that a similarity value is calculated for each cluster, and two clusters are combined if, and only if, both clusters have a similarity value greater than a predetermined threshold (i.e. 0.5). If either cluster fails to have a similarity value greater than 0.5, then neither cluster is combined. This is in direct conflict with the limitation of claim 2, which required that the number of common files be considered "relative to" the total number of documents in all the clusters under consideration.

To remove any unintended ambiguity, this limitation of claim 2 was rephrased in more exact terms and moved into its base claim 1. Claim 1 is thus amended to more clearly define the determination of the degree of relationship between multiple clusters under consideration as

"the number of distinct files common to all of said clusters under consideration multiplied by a predefined multiplication factor divided by a total sum of all the files in said clusters under consideration".

Support for this limitation is found at least in page 14, first paragraph of the present Patent Application.

Claim 2 is subsequently amended to specify that the multiplication factor is equal to the number of clusters under consideration. This limitation is not taught or suggested in any of the cited prior art references. Also, new claim 14 adds the further limitation of limiting the number of clusters under consideration to two.

Independent claims 10 and 12 were amended to include a similar clarification of claim limitations as in independent claim 1. Similarly, new claims 16 and 17, which depend from claim 10, and new claims 18 and 19, which depend from claim 12, introduce similar limitations as claims 2 and 14, respectively.

In reference to the second aspect relating to the use of the terms "AND" and "OR" in claim 7, it would appear that the Office Action has interpreted the use of "AND" and "OR" as user-submitted, Boolean filtering options in a database search engine. This is incorrect. The terms "AND" and "OR" are used in the specification of the present application as labels to define two types of listing formats (i.e. naming formats for merged clusters). By reviewing the combination of "AND" and "OR" listing formats in the name structure of a combined cluster (consisting of multiple merged clusters), one can determine the relative similarity between merged clusters without having to expand the clusters.

The Office Action asserts that, "it would have been obvious to one of ordinary skill in the art at the time of the invention to provide displays based on the relationships between clusters...providing the benefit of having simplified the understanding of the search results for the user". Applicants respectfully disagree. Firstly, claim 7 did not merely recite a naming scheme that provided a general "relationships between clusters", as is assumed by the Office Action. Indeed, the mere concatenation of combined clusters would provide a "relationships between clusters" by indicating which clusters were merged.

Rather, claim 7 required a naming scheme that would indicate, "the degree of relation among said clusters merged together." That is, claim 7 requires a determination of relative levels of similarity between merged clusters,

and to indicate these relative levels in the naming format of the merged clusters. Zamir provides no method by which to categorize levels of similarity between merged clusters, and thus there is no incentive to introduce a naming scheme to convey information regarding levels of similarity between merged clusters.

Nonetheless, to remove any unintended ambiguity, claim 7 is rewritten in independent form incorporating limitations from its previously base claim 6, and is further amended to : (1) replace the labels "AND" and "OR" with terms "first listing format" and "second listing format"; and (2) to more clearly define multiple levels of similarity by expressing a first range of similarity between first and second predetermined values and a second range of similarity between second and third predetermined values. Support for recitation (2) is found at least in page 30, line 1 to page 31, line 6 of the present Patent Application.

Amended Claim 8 defines the first and second listing formats. New claim 15 identifies the first and second listing formats as "AND" and "OR" listing formats, respectively.

Independent claims 11 and 13 are amended to include a similar clarification of claim limitations as in independent claim 7. Similarly, new claims 20 and 21, which depend from claim 11, and new claims 22 and 23, which depend from claim 13, introduce similar limitations as claims 8 and 15, respectively.

In view of the foregoing amendments and remarks, Applicants respectfully request favorable reconsideration of the present application.

Respectfully submitted,

Rosalio Haro
Registration No. 42,633


Please address all correspondence to:

Epson Research and Development, Inc.
Intellectual Property Department
150 River Oaks Parkway, Suite 225
San Jose, CA 95134
Phone: (408) 952-6000
Facsimile: (408) 954-9058
Customer No. 20178

Date: August 9, 2004